



(Cahiers Mathématiques de l'Université de Sherbrooke)

Titre : Les qualifications des contraintes en optimisation à contraintes non linéaires

Auteur(s) : Luc Marchand

Revue : CaMUS (Cahiers Mathématiques de l'Université de Sherbrooke)

Volume : 5

Année : 2014

Pages : 23-36

Éditeur : Université de Sherbrooke. Département de Mathématiques

URI : Repéré à : <http://camus.math.usherbrooke.ca/revue.html>

Page vide laissée intentionnellement

Les qualifications des contraintes en optimisation à contraintes non linéaires

Luc Marchand

RÉSUMÉ Cet article consiste en une vulgarisation du concept de qualification des contraintes du premier ordre dans un programme d'optimisation non linéaire. Des exemples de problèmes explicitant les différences entre plusieurs qualifications sont énoncés, permettant de mieux visualiser ce qu'impliquent ces qualifications. Les relations entre trois qualifications sont approfondies, c'est-à-dire l'indépendance linéaire (LICQ), la qualification de Mangasarian-Fromovitz (MFCQ) et la condition du rang constant (CRCQ). D'autres qualifications de nature plus analytique sont ensuite présentées.

1 Introduction

1.1 Mise en place du problème d'optimisation

Soit le programme non linéaire d'optimisation :

$$\left\{ \begin{array}{ll} \min & z = f(x) \\ \text{s.à.} & g_i(x) \leq 0 \quad i = 1, \dots, m \\ & h_j(x) = 0 \quad j = 1, \dots, p \end{array} \right. \quad (\text{NLP})$$

où $f(x)$, $g_i(x)$ et $h_j(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ sont des fonctions différentiables. La notation NLP est un acronyme pour *Non Linear Program*. Nous nous limitons à trouver un minimum local x^* du programme (NLP). Le point de départ de l'étude de tels problèmes est d'établir des conditions nécessaires d'optimalité, c'est-à-dire se poser la question : «Que peut-on déduire du fait que x^* est un minimum local ? ».

Pour poursuivre, nous introduisons la fonction

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^p \mu_j h_j(x)$$

avec $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^p$. Cette fonction est appelée le lagrangien associé au problème (NLP). Les scalaires λ_i et μ_i sont appelés les multiplicateurs de Lagrange.

Remerciements particuliers au professeur Jean-Pierre Dussault et au comité de révision pour leurs conseils pour la rédaction de ce document.

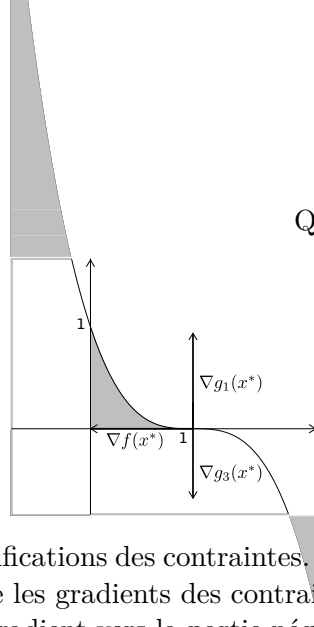


FIGURE 1 – Échec des qualifications des contraintes. Le point optimal étant situé en $(1, 0)^t$, on remarque que les gradients des contraintes génèrent l'axe vertical. Notre fonction ayant son gradient vers la partie négative de l'axe horizontal, on remarque que le résultat de KKT n'est pas vérifié.

En optimisation non linéaire avec contraintes, les conditions nécessaires de Karush-Kuhn-Tucker (conditions de KKT) sont les plus utilisées. Celles-ci stipulent qu'en un point stationnaire du Lagrangien \bar{x} , il existe des vecteurs $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p$ tels que :

$$\begin{aligned} \nabla \mathcal{L}(\bar{x}, \lambda, \mu) &= \nabla f(\bar{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\bar{x}) + \sum_{j=1}^p \mu_j \nabla h_j(\bar{x}) = 0 \\ \lambda_i &\geq 0, \lambda_i g_i(\bar{x}) = 0 \quad i = 1, \dots, m, j = 1, \dots, p \end{aligned} \quad (1)$$

Nous noterons par I l'ensemble des indices des contraintes $g_i(x)$. Donc la cardinalité de I est m . Nous noterons aussi par I^* l'ensemble des indices de I tels que $g_i(\bar{x}) = 0$. Les contraintes dont l'indice est dans I^* seront appelées contraintes actives au point \bar{x} et seuls leurs multiplicateurs seront éventuellement non-nuls.

Intuitivement, nos conditions nécessaires (1) signifient que le gradient de la fonction objectif peut être reconstruit par une combinaison linéaire des gradients des contraintes actives.

On désire que les points optimaux de $f(x)$ soient des points stationnaires du lagrangien. L'exemple suivant nous montrera qu'il faut une hypothèse supplémentaire afin que le résultat soit vérifié en un optimum local.

1.2 Exemple de non existence de multiplicateurs

Exemple 1.1. Soit le programme non linéaire (voir Figure 1) :

$$\left\{ \begin{array}{ll} \min & z = -x_1 \\ \text{s.à.} & g_1(x) = (x_1 - 1)^3 + x_2 \leq 0 \\ & g_2(x) = -x_1 \leq 0 \\ & g_3(x) = -x_2 \leq 0 \end{array} \right.$$

Le point optimal est $x^* = (1, 0)^t$. En x^* , on remarque que les contraintes $g_1(x)$ et $g_3(x)$ sont satisfaites avec égalité. On en déduira donc que $I^* = \{1, 3\}$ alors que $I = \{1, 2, 3\}$ puisqu'il y a trois contraintes d'inégalité. Les multiplicateurs

λ_i pouvant être non nuls sont donc λ_1 et λ_3 . Les conditions (1) s'écrivent donc (après avoir géré les équations $\lambda_i g_i(\bar{x}) = 0$ et ainsi, retiré $g_2(x)$ puisque $\lambda_2 = 0$) :

$$\begin{aligned}\nabla \mathcal{L}(\bar{x}, \lambda, \mu) &= \nabla f(\bar{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\bar{x}) = 0 \\ &= \begin{pmatrix} -1 & 0 \end{pmatrix} + \begin{pmatrix} \lambda_1 & \lambda_3 \end{pmatrix} \begin{pmatrix} 3(\bar{x}_1 - 1)^2 & 1 \\ 0 & -1 \end{pmatrix} = 0\end{aligned}$$

et en x^* nous avons

$$\nabla \mathcal{L}(x^*, \lambda, \mu) = \begin{pmatrix} -1 & 0 \end{pmatrix} + \begin{pmatrix} \lambda_1 & \lambda_3 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$

On remarque qu'il n'existe pas de solution λ^* satisfaisant cette équation et, malgré cela, x^* est bel et bien un minimum local, même global.

Remarque 1.2. Dans le cas linéaire, on ne peut trouver d'exemple comme ci-dessus. En fait, l'existence des multiplicateurs sous contraintes linéaires en un minimum local ne nécessite aucune hypothèse supplémentaire.

Ainsi, on doit exiger une hypothèse supplémentaire en un optimum x^* pour que celui-ci soit un point stationnaire du lagrangien (1). Cette hypothèse est qu'une qualification des contraintes soit respectée en \bar{x} . Mais qu'est-ce qu'une qualification des contraintes ?

2 Qualifications des contraintes d'ordre 1

Cette section constitue le corps principal de l'article. Nous y présentons la définition d'une qualification des contraintes (que j'acronymerai QC) au sens large, suivi des définitions de quelques QC. Plusieurs exemples seront aussi énoncés ici.

J'indique ici que les exemples qui permettront d'illustrer les différentes conditions présentées sont principalement composés de contraintes d'inégalités. Ceci est fait ainsi afin qu'il soit plus aisé de voir les subtilités entre les qualifications et réduire autant que possible la dimension des problèmes pour les illustrer plus facilement.

2.1 Définition et arborescence

L'exemple 1.1 montre que, même en un minimum local x^* , il n'existe pas nécessairement de multiplicateurs de Lagrange $\lambda^* \geq 0$ et μ^* libre tels que $\nabla f(x^*) + \lambda^* \nabla g(x^*) + \mu^* \nabla h(x^*) = 0$. Une qualification des contraintes est simplement une hypothèse suffisante à l'existence de ces multiplicateurs en x^* pour que l'équation (1) soit satisfaite. Intuitivement, on peut voir une QC comme une hypothèse nous certifiant que la linéarisation des contraintes $\nabla g(x)$ et $\nabla h(x)$ agit localement de manière similaire aux contraintes $g(x)$ et $h(x)$.

Remarque 2.1. La linéarisation d'une fonction $g(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ en x_0 consiste en l'hyperplan de dimension n passant par x_0 et ayant les mêmes dérivées directionnelles que la fonction $g(x)$ au point x_0 . Formellement, la linéarisation $l_{x_0}(x)$ en x_0 d'une fonction $g(x)$ est donnée par $l_{x_0}(x) = g(x_0) + \nabla g(x_0)(x - x_0)$. Cette fonction évaluée en un point \bar{x} suffisamment près de x_0 devrait donner une bonne approximation de $g(\bar{x})$. Aussi, la linéarisation de $g(x)$ en x_0 est l'hyperplan tangent à la fonction $g(x)$ en x_0 . Les propriétés de cette linéarisation dépendent grandement du vecteur $\nabla g(x_0)$. Bien que ce soit un abus de notation, nous appelons $\nabla g(x_0)$ la linéarisation de $g(x)$ en x_0 pour simplifier et alléger le discours.

En revenant à l'exemple 1.1, on remarque que la linéarisation des contraintes actives

$$\nabla g_{I^*}(x^*) = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$

nous laisse croire que l'axe des x_1 reste réalisable en x^* , donc que l'on pourrait se déplacer autant vers la gauche que vers la droite à partir de $(1, 0)^t$ en restant réalisable. Toutefois, le vrai domaine réalisable de ce problème se situe entièrement à gauche de $(1, 0)^t$.

Il existe évidemment plusieurs qualifications différentes pour lesquelles les multiplicateurs de Lagrange ont certaines propriétés. Dans ce qui suit, on définira quatre de ces qualifications et leurs relations entre elles. La plupart des qualifications des contraintes sont notées par un acronyme tiré de l'anglais. Afin de familiariser les lecteurs avec la terminologie utilisée dans le domaine, j'utiliserai aussi ces acronymes anglais en expliquant leur provenance.

2.2 LICQ

Définition 2.2. La qualification des contraintes d'indépendance linéaire LICQ (*Linear Independence Constraints Qualification*) est satisfaite en un point x^* lorsque la matrice

$$\begin{pmatrix} \nabla h(x^*) \\ \nabla g_{I^*}(x^*) \end{pmatrix}$$

est de plein rang ligne, c'est-à-dire que le rang de cette matrice égale le nombre de lignes de la matrice.

Le théorème suivant nous montre que LICQ est bel et bien une QC.

Théorème 2.3. [FM90] Soit x^* un minimum local de (NLP). Si LICQ est satisfait en x^* , alors il existe un vecteur $\lambda^* \in \mathbb{R}^m$ et $\mu^* \in \mathbb{R}^p$ tel que $\nabla f(x^*) + \lambda^* \nabla g(x^*) + \mu^* \nabla h(x^*) = 0$, $\lambda_i^* g_i(x^*) = 0$ et $\lambda^* \geq 0$. De plus, ces vecteurs sont uniques. \square

En pratique, certaines méthodes de recherche de solutions de problèmes de la forme (NLP) convergent plus rapidement lorsque x^* est un point satisfaisant LICQ. C'est pourquoi il s'agit d'une des qualifications des contraintes les plus

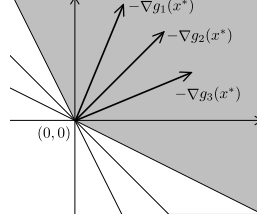


FIGURE 2 – Échec de LICQ. Il est évident, au point optimal $(0,0)^t$, que les gradients des contraintes actives ne peuvent pas être linéairement indépendants, puisque le nombre de contraintes actives en ce point est plus grand que la dimension du problème.

utilisées et connues. On remarque toutefois que LICQ est une qualification très forte, comme le montre l'exemple suivant.

Exemple 2.4. Soit le programme d'optimisation à contraintes linéaires (voir Figure 2) :

$$\begin{cases} \min & z = (x_1 + 1)^2 + (x_2 + 1)^2 \\ \text{s.à.} & g_1(x) = -0,5x_1 - x_2 \leq 0 \\ & g_2(x) = -x_1 - x_2 \leq 0 \\ & g_3(x) = -2x_1 - x_2 \leq 0 \end{cases}$$

La solution de ce problème est $x^* = (0,0)^t$. Or, en x^* , trois contraintes sont actives dans un espace de dimension 2, ce qui implique que la matrice

$$\nabla g_{I^*}(x^*) = \begin{pmatrix} -0,5 & -1 \\ -1 & -1 \\ -2 & -1 \end{pmatrix}$$

ne peut être de plein rang ligne, ainsi LICQ n'est pas satisfaite.

On remarque aussi qu'il existe plusieurs vecteurs de multiplicateurs qui satisfont aux conditions (1). En particulier, puisque $\nabla f(0) = \begin{pmatrix} 2 & 2 \end{pmatrix}$, on obtient le système

$$\begin{aligned} 2 - 0,5\lambda_1 - 1\lambda_2 - 2\lambda_3 &= 0 \\ 2 - \lambda_1 - \lambda_2 - \lambda_3 &= 0 \\ \lambda_i &\geq 0 \quad (i = 1, 2, 3) \end{aligned}$$

qui admet les solutions $0 \leq \lambda_1 \leq \frac{4}{3}$, $\lambda_2 = 2 - \frac{3\lambda_1}{2}$ et $\lambda_3 = \frac{\lambda_1}{2}$.

Nous venons de voir que LICQ peut échouer même dans un programme sous contraintes linéaires qui, comme cité à la remarque 1.2, ne nécessite pas de QC. En effet, dans le problème précédent, puisqu'il y a plus d'un multiplicateur λ^* vérifiant les conditions d'optimalité de KKT, on pouvait s'attendre à ce que LICQ échoue.

2.3 MFCQ

Définition 2.5. La qualification des contraintes de Mangasarian-Fromovitz MFCQ est satisfaite en un point x^* lorsque la matrice $\nabla h(x^*)$ est de plein rang ligne et qu'il existe un vecteur $v \in \mathbb{R}^m$ tel que $\nabla g_{I^*}(x^*)v < 0$ et $\nabla h(x^*)v = 0$.

Cette qualification signifie simplement qu'il existe un vecteur qui entre dans l'intérieur du domaine réalisable sans rester sur la frontière des contraintes d'inégalité actives tout en restant sur les contraintes d'égalités.

Cette QC est aussi très connue, ce qui est expliqué par le théorème suivant.

Théorème 2.6. [FM90] Soit x^* un minimum local de la fonction $f(x)$ du problème (NLP). On a que x^* satisfait MFCQ si et seulement si l'ensemble Λ^* des vecteurs satisfaisant $\lambda^* \in \mathbb{R}^m, \mu^* \in \mathbb{R}^p$ tels que $\lambda^* \geq 0, \nabla f(x^*) + \lambda^* \nabla g(x^*) + \mu^* \nabla h(x^*) = 0$ et $\lambda_i^* g_i(x^*) = 0$ est non vide et borné. \square

Puisque plusieurs algorithmes d'optimisation non linéaire utilisent les multiplicateurs de Lagrange comme outil de recherche de solutions, le fait que l'ensemble des multiplicateurs soit borné sous MFCQ est d'une grande utilité. De plus, tout comme LICQ, cette QC est plutôt facile à vérifier. La suffisance de la condition du théorème précédent est aussi remarquable.

Corollaire 2.7. LICQ implique MFCQ.

Démonstration. Allons-y par la contraposée. On suppose MFCQ non satisfaite. Donc l'ensemble des multiplicateurs de Lagrange est soit vide, soit non borné. Dans ces deux cas, l'ensemble des multiplicateurs n'est pas un singleton, et LICQ ne peut être satisfaite. \square

Remarque 2.8. L'exemple 2.4 satisfait MFCQ sans satisfaire LICQ. Les multiplicateurs de ce problème respectent $0,5\lambda_1 = \lambda_3, 2 - 1,5\lambda_1 = \lambda_2$. La condition $\lambda_i \geq 0$ nécessite donc $\lambda_1 \in [0, \frac{4}{3}]$. On remarque ainsi que l'ensemble

$$\Lambda^* = \left\{ (\lambda_1, \lambda_2, \lambda_3) \mid 0 \leq \lambda_1 \leq \frac{4}{3}, \lambda_2 = 2 - \frac{3\lambda_1}{2}, \lambda_3 = \frac{\lambda_1}{2} \right\}$$

des multiplicateurs de Lagrange possibles dans les équations (1) est de cardinalité infinie et est borné.

Une condition équivalente à MFCQ est donnée par l'énoncé suivant :

Proposition 2.9. MFCQ est satisfaite en x^* si et seulement si la condition suivante est satisfaite :

$$\left\{ \begin{array}{l} \lambda \nabla g(x^*) + \mu \nabla h(x^*) = 0 \\ \lambda g(x^*) = 0 \\ \lambda \geq 0 \end{array} \right\} \implies \left\{ \begin{array}{l} \lambda = 0 \\ \mu = 0 \end{array} \right\} \quad (2)$$

\square

La démonstration de cette proposition utilise le théorème de dualité faible de programmation linéaire. Il s'agit d'une manière intéressante de démontrer des résultats à partir de théorèmes d'alternative (dont fait partie le théorème de dualité faible). La preuve est donc incluse à l'annexe A de ce document.

L'exemple suivant montre une situation telle que, pour toute fonction objectif deux fois continuellement différentiable admettant $(0,0)^t$ comme minimum, il existe des multiplicateurs sans que MFCQ soit satisfaite. Il est donc possible d'affaiblir davantage MFCQ, toutefois nous perdrons le fait que l'ensemble des multiplicateurs soit borné.

Exemple 2.10. Soit le programme d'optimisation à contraintes non linéaires (voir Figure 3) :

$$\left\{ \begin{array}{ll} \min & z = (x_1 + 1)^2 + x_2^2 \\ \text{s.à.} & g_1(x) = x_1^2 - x_2 \leq 0 \\ & g_2(x) = -x_1^2 + x_2 \leq 0 \\ & g_3(x) = -x_1 \leq 0 \end{array} \right.$$

Le domaine réalisable de ce programme est la branche $x \geq 0$ de la parabole. La solution de ce programme de minimisation est située au point $x^* = (0,0)^t$. On remarque ici que MFCQ n'est pas satisfaite puisqu'il n'y a pas d'intérieur strict des contraintes.

Si l'on calcule le lagrangien de ce problème en x^* , nous obtenons

$$\begin{pmatrix} 2 & 0 \end{pmatrix} + \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 0 & 1 \\ -1 & 0 \end{pmatrix} = 0$$

On observe donc que $\lambda_1 = \lambda_2$ et $\lambda_3 = 2$ satisfont ces équations pour toute valeur de $\lambda_1 \in \mathbb{R}$ et $\lambda_1 \geq 0$. Il y a donc un ensemble non borné de multiplicateurs possibles puisque les gradients des contraintes, $\nabla g_1(x)$ et $\nabla g_2(x)$, sont linéairement dépendants et de directions opposées pour tout point réalisable.

MFCQ réagit très mal aux contraintes d'égalité qui se dupliquent en deux contraintes d'inégalité. C'est le cas de l'exemple précédent avec les contraintes $g_1(x)$ et $g_2(x)$. En effet, si l'on définit $h_1(x) = x_1^2 - x_2 = 0$, l'ensemble réalisable reste le même en prenant les contraintes $h_1(x)$ et $g_3(x)$ et, avec une telle définition des contraintes, MFCQ est satisfaite en x^* .

2.4 CRCQ

Définition 2.11. La qualification des contraintes du rang constant CRCQ (*Constant Rank Constraints Qualification*) est satisfaite en un point x^* lorsqu'il existe un voisinage $X \subset E$ autour de x^* tel que pour tout sous-ensemble $K \subset I^*$ et $K' \subset \{1, \dots, p\}$, le rang de la matrice

$$\begin{pmatrix} \nabla h_j(\bar{x}) \\ \nabla g_i(\bar{x}) \end{pmatrix} \quad (i \in K, j \in K')$$

est constant pour tout $\bar{x} \in X$.

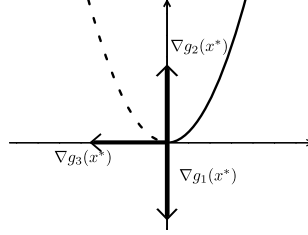


FIGURE 3 – Échec de MFCQ. Le domaine réalisable est la branche pleine de la parabole. La problématique vient du fait qu'il n'existe pas d'intérieur du domaine réalisable régi par les inégalités. Une forme alternative des contraintes (si l'on considère $g_1(x)$ et $g_2(x)$ comme une seule contrainte d'égalité) résout cette problématique.

Il s'agit bel et bien d'une qualification des contraintes (donc assurant l'existence des multiplicateurs de Lagrange). Toutefois, elle a été développée beaucoup plus tard que les autres qualifications présentées dans cet article. Cette qualification a été amenée principalement par Spingarn [Spi83] et Janin [Jan84]. Nous avons évidemment l'implication suivante :

Proposition 2.12. *LICQ implique CRCQ.*

Démonstration. LICQ satisfaite signifie que la matrice $\nabla g_i(x^*), i \in I^*$ est de plein rang ligne. Donc aucune ligne de la matrice n'est linéairement dépendante des autres. Puisque les fonctions $g_i(x)$ sont deux fois continuellement différentiable, alors $\nabla g_i(x)$ sont continues, donc il existe un voisinage X de x^* tel que la matrice $\nabla g_i(\bar{x}), \bar{x} \in X$ est aussi de plein rang ligne, et on a CRCQ.

□

Remarque 2.13. Le fait que les contraintes $g_i(x), i \in I^*$ soient linéaires est suffisant à ce que CRCQ soit satisfaite en x^* puisque la matrice composée des lignes $\nabla g_i(x), i \in I^*$ est constante.

Les relations entre CRCQ et MFCQ sont plus subtiles. D'abord, remarquons que CRCQ est satisfaite pour l'exemple 2.10, donc CRCQ n'implique visiblement pas MFCQ. En effet, pour tout point de \mathbb{R}^2 (donc en particulier pour un voisinage de x^*), les gradients $\nabla g_1(x)$ et $\nabla g_2(x)$ sont linéairement dépendants (matrice de rang 1). De plus, les gradients $\nabla g_1(x)$ et $\nabla g_3(x)$ (ou $\nabla g_2(x)$ et $\nabla g_3(x)$) sont linéairement indépendants (de rang 2). Donc pour tout sous-ensemble des contraintes actives, la matrice des gradients des dites contraintes est de rang constant dans \mathbb{R}^2 .

L'exemple suivant montre que MFCQ n'implique pas CRCQ non plus.

Exemple 2.14. Soit le programme d'optimisation à contraintes non linéaires (voir Figure 4) :

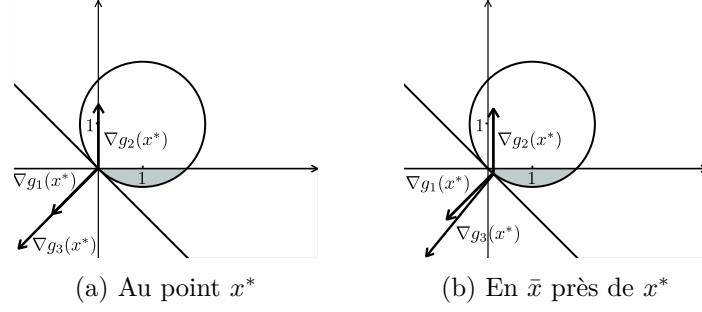


FIGURE 4 – Échec de CRCQ. Le point optimal x^* est en $(0, 0)^t$. On remarque que le long de la droite $x_2 = x_1$, les gradients des contraintes 1 et 3 sont linéairement dépendants (Figure 4a), mais pas ailleurs (Figure 4b). Puisque $x_2 = x_1$ passe par x^* , on a qu'il n'existe pas de voisinage autour de x^* de rang constant.

$$\begin{cases} \min & z = (x_1)^2 + (x_2 - 1)^2 \\ \text{s.à.} & g_1(x) = -x_1 - x_2 \leq 0 \\ & g_2(x) = x_2 \leq 0 \\ & g_3(x) = (x_1 - 1)^2 + (x_2 - 1)^2 \leq 0 \end{cases}$$

L'ensemble réalisable est borné par l'arc du cercle défini par $g_3(x) = 0$ et sous l'axe des x_1 . Le minimum local de ce problème est situé en $x^* = (0, 0)^t$ où les trois contraintes sont actives. En ce point, le vecteur $v = (1, -\frac{1}{4})^t$ nous amène dans l'intérieur du domaine réalisable, donc MFCQ est satisfaite. En effet, on voit que

$$\begin{pmatrix} \nabla g_1(x^*) \\ \nabla g_2(x^*) \\ \nabla g_3(x^*) \end{pmatrix} v = \begin{pmatrix} -1 & -1 \\ 0 & 1 \\ -2 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ -1/4 \end{pmatrix} = \begin{pmatrix} -3/4 \\ -1/4 \\ -3/2 \end{pmatrix}$$

et ainsi $\nabla g_i(x^*)v < 0, \forall i \in I^*$, d'où MFCQ satisfaite.

Toutefois, si l'on prend l'ensemble d'indices $K = \{1, 3\}$, alors la matrice des $\nabla g_i(x^*)$ pour $i \in K$ s'écrit

$$\begin{pmatrix} -1 & -1 \\ 2x_1 - 2 & 2x_2 - 2 \end{pmatrix}$$

n'admet pas de voisinage de rang constant autour de x^* (la matrice est de rang 1 sur la droite $x_1 = x_2$ et de rang 2 ailleurs). Ainsi, CRCQ n'est pas satisfaite en x^* .

Remarque 2.15. Lu [Lu11] a récemment prouvé que, dans un problème non paramétrique (tels que ceux présentés plus haut), CRCQ implique MFCQ après modification de la forme des contraintes. On peut voir cela dans l'exemple 2.10, si l'on redéfinit $g_1(x)$ et $g_2(x)$ par une seule contrainte d'égalité $h_1(x)$. Les modifications apportées ne doivent évidemment pas modifier le domaine réalisable.

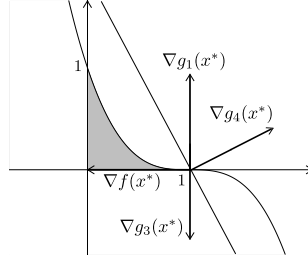


FIGURE 5 – Échec de MFCQ et CRCQ. MFCQ échoue car le seul vecteur réalisable est dans la direction $(-1, 0)^t$ qui n'est pas strictement réalisable ($g_3(x)$ reste satisfaite avec égalité le long de ce vecteur). De plus, CRCQ n'est pas satisfaite puisque pour les contraintes $g_1(x)$ et $g_3(x)$, il n'existe pas de voisinage de rang constant en x^* .

L'exemple suivant venant de [CW05] nous servira à déduire une autre qualification des contraintes historiquement importante.

Exemple 2.16. Soit le programme d'optimisation non linéaire (voir Figure 5) :

$$\left\{ \begin{array}{ll} \min & z = -x_1 \\ \text{s.à.} & g_1(x) = (x_1 - 1)^3 + x_2 \leq 0 \\ & g_2(x) = -x_1 \leq 0 \\ & g_3(x) = -x_2 \leq 0 \\ & g_4(x) = 2x_1 + x_2 - 2 \leq 0 \end{array} \right.$$

Il s'agit de l'exemple 1.1 auquel on a ajouté une contrainte active en $x^* = (1, 0)^t$. On remarque que MFCQ n'est pas satisfaite (les seules directions réalisables sont le long de l'axe des x_1 , pas à l'intérieur strict du domaine réalisable) et CRCQ non plus (l'ensemble d'indices $K = \{1, 3\}$ n'est pas de rang constant dans tout voisinage de x^*). On remarque toutefois qu'il existera des multiplicateurs de Lagrange (λ^*, μ^*) . En effet, on trouve facilement que $\nabla f(x^*) = (-1, 0)$ et que $\nabla g_1(x^*) = (0, 1)$, $\nabla g_3(x^*) = (0, -1)$ et $\nabla g_4(x^*) = (2, 1)$. En prenant $\lambda_4 = 1/2$ et $\lambda_1 - \lambda_3 = -1/2$, on obtient les équations de KKT satisfaites, justifiant ainsi qu'il existe des multiplicateurs de Lagrange pour ce problème.

Remarque 2.17. Cet exemple nous permet d'observer quelque chose de remarquable en soi. L'ajout de la contrainte $g_4(x)$ ne modifie d'aucune manière le domaine réalisable. On pourrait même affirmer qu'elle semble inutile au problème. Toutefois, son ajout nous permet de travailler sur les multiplicateurs de Lagrange dans la recherche du point optimal. On doit toutefois observer que son ajout ne nous permet pas nécessairement d'avoir une qualification forte telle que LICQ. En regardant les choses sous un angle différent, on peut aussi en déduire que les qualifications des contraintes ne sont pas des conditions sur la structure du domaine réalisable, mais vraiment des conditions sur la structure des contraintes. En fait, considérant l'ensemble généré par les gradients des contraintes actives, on aura qu'il existera des multiplicateurs pour toute fonction admettant x^* comme minimum local.

L'exemple précédent nous motive à présenter succinctement d'autres qualifications des contraintes. Afin de faire une présentation complète, on ne peut passer ces qualifications sous silence. Nous ne présentons toutefois pas les détails de ces qualifications puisque celles-ci débordent de l'objectif de l'article : présenter des qualifications vérifiables par des méthodes issues de l'algèbre linéaire.

2.5 Qualifications des contraintes de nature analytique

Les qualifications des contraintes vues précédemment s'expriment à l'aide du langage de l'algèbre linéaire : l'indépendance linéaire de vecteurs et le rang de matrices. Les qualifications suivantes s'expriment à l'aide de notions topologiques et analytiques. Ainsi, je ne ferai qu'une brève présentation des définitions en mettant quelques résultats plus importants en référence.

Définition 2.18. La qualification des contraintes de Kuhn-Tucker (KTCQ) est satisfaite en x^* si, pour tout vecteur non nul z tel que $\nabla g_i(x^*)z \leq 0$ (pour tout $i \in I^*$ et $\nabla h_j(x^*)z = 0, j = 1, \dots, p$), z est tangent à un arc $\alpha(\theta)$ une fois différentiable tel que $\alpha(0) = x^*$ et est contenu dans le domaine réalisable.

Il s'agit de la première qualification des contraintes à avoir été publiée, d'abord par Karush dans [Kar39], puis par Kuhn et Tucker dans [KT51]¹.

Cette QC, bien qu'importante historiquement, est maintenant considérée de moins grande importance avec la découverte des qualifications des contraintes d'Abadie et de Guignard, des qualifications encore plus faibles que KTCQ (ces deux qualifications seront présentées plus loin).

De plus, KTCQ est difficile à vérifier, limitant son utilisation par des algorithmes.

Remarque 2.19. Dans l'exemple 1.1, KTCQ n'est pas satisfaite. En effet, il existe deux vecteurs non nuls z_i tels que $\nabla g_i(x^*)z_i \leq 0 (i = 1, 2)$; il s'agit des vecteurs $z_1 = \alpha(1, 0)^t$ et $z_2 = \beta(-1, 0)^t$, $\alpha, \beta \in \mathbb{R}, \alpha, \beta \geq 0$. Or, pour le vecteur z_1 , il n'existe pas d'arc différentiable tel que KTCQ soit satisfaite.

Remarque 2.20. Dans l'exemple 2.16, KTCQ est satisfaite. Ceci est expliqué par le fait que le vecteur problématique z_1 n'est plus tel que $\nabla g_i(x^*)z_1 \leq 0$ (car $\nabla g_4(x^*)z_1 > 0$). KTCQ est donc désormais vérifiée.

Théorème 2.21. *CRCQ implique KTCQ.*

Démonstration. Voir [Jan84, Proposition 2.3]

□

La qualification des contraintes suivante utilise les définitions de cône tangent aux contraintes et de cône linéarisé. Plusieurs articles incluent ces définitions (par exemple [Bur12], [Jan84], [Fle87] ou [Sol11]).

¹Karush a découvert cette qualification (et les conditions de KKT) 12 ans avant Kuhn et Tucker, au cours de sa maîtrise. Voir [Cot12] pour davantage d'informations.

Définition 2.22. [Bur12] La qualification des contraintes d’Abadie (ACQ) est satisfaite en x^* si le cône tangent aux contraintes actives est le même que le cône linéarisé au point x^* :

$$T(x^*) = L(x^*)$$

Puis, finalement, la qualification de Guignard nécessite un peu d’analyse fonctionnelle, en particulier la définition du cône polaire d’un ensemble convexe, qui ne sera pas présentée ici (voir [Roc70] pour plus d’informations).

Définition 2.23. [Bur12] La qualification des contraintes de Guignard (GCQ) est satisfaite en x^* si, avec les mêmes notations qu’à la définition 2.22, on a

$$T(x^*)^\circ = L(x^*)^\circ$$

où E° représente le polaire de l’ensemble E .

3 Conclusion

Les multiplicateurs de Lagrange étant un outil très puissant en optimisation non linéaire, il est naturel d’étudier les qualifications des contraintes permettant d’assurer l’existence de ces multiplicateurs en un point optimal. Dans cet article, nous expliquons certaines de ces qualifications et une intuition de leur comportement dans différentes situations. Nous énonçons aussi quelques exemples permettant de visualiser ces qualifications des contraintes et quand celles-ci peuvent échouer.

Il y a encore beaucoup à faire sur les qualifications des contraintes, entre autres au sujet des qualifications des contraintes d’ordre 2, non présentées dans cet article. CRCQ est aussi actuellement un sujet de recherche dans la communauté.

Références

- [Bur12] J. BURKE : Constraint Qualifications in Nonlinear Programming. Notes de cours Math 516, Printemps 2012.
- [Cot12] R. W. COTTLE : William Karush and the KKT theorem. *Doc. Math., J. DMV*, pages 255–269, 2012.
- [CW05] A.C. CHIANG et K. WAINWRIGHT : *Fundamental methods of mathematical economics*. McGraw-Hill, 2005.
- [Fle87] R. FLETCHER : *Practical Methods of Optimization*. Wiley-Interscience, Chichester, 1987.

- [FM90] A. V. FIACCO et G. P. MCCORMICK : *Nonlinear Programming : Sequential Unconstrained Minimization Techniques*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1990.
- [Jan84] R. JANIN : Directional Derivative of the Marginal Function in Nonlinear Programming. *Math. Program. Study*, 21:110–126, 1984.
- [Kar39] W. KARUSH : Minima of Functions of Several Variables with Inequalities as Side Conditions. Mémoire de maîtrise, Université de Chicago, 1939.
- [KT51] H. W. KUHN et A. W. TUCKER : Nonlinear Programming. *In Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481–492. University of California Press, 1951.
- [Lu11] S. LU : Implications of the Constant Rank Constraint Qualification. *Math. Prog.*, 126(2):365–392, 2011.
- [Roc70] R. Tyrrell ROCKAFELLAR : *Convex analysis*. Princeton Mathematical Series. Princeton University Press, Princeton, N. J., 1970.
- [Sol11] M.V. SOLODOV : Constraint Qualifications. *Wiley Encyclopedia of Operations Research and Management Science*, pages 255–269, 2011.
- [Spi83] J. E. SPINGARN : Partial inverse of a monotone operator. *Applied Mathematics and Optimization*, 10(1):247–265, 1983.

LUC MARCHAND

DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE SHERBROOKE

Courriel: luc.marchand@usherbrooke.ca

A Démonstration de la proposition 2.9

Démonstration. Cette démonstration vient du professeur Jean-Pierre Dussault.

Notons q la cardinalité de I^* .

Soit le programme linéaire

$$\begin{array}{ll}
 \min_{(d,u) \in \mathbb{R}^n \times \mathbb{R}} & z = u \\
 \text{s. à.} & \nabla g_{I^*}(x^*)d \leq ue \\
 & \nabla h(x^*)d = 0
 \end{array} \tag{3}$$

où $e = (1, 1, \dots, 1)^t \in \mathbb{R}^q$. On observe que $(d, u) = (0, 0)^t$ est toujours réalisable. La condition $\lambda g(x^*) = 0$ est camouflée dans le fait que l'on ne travaille que sur les contraintes d'indice dans I^* . Le dual de ce problème est

$$\begin{aligned}
& \max_{(\lambda, \mu) \in \mathbb{R}^q \times \mathbb{R}^p} && 0 \\
& \text{s. à.} && \lambda \nabla g_{I^*}(x^*) + \mu \nabla h(x^*) = 0 \\
& && \lambda e = 1 \\
& && \lambda \geq 0
\end{aligned} \tag{4}$$

(\Rightarrow) On suppose d'abord que MFCQ est satisfaite. Alors, il existe une direction $\tilde{d} \in \mathbb{R}^n$ telle que

$$\begin{aligned}
& \nabla g_{I^*}(x^*) \tilde{d} < 0 \\
& h(x^*) \tilde{d} = 0
\end{aligned}$$

Posons $\nabla g_{I^*}(x^*) \tilde{d} \leq \alpha e < 0$. Ainsi, le problème primal (3) est tel que

$$\begin{aligned}
& \min_{(d, u) \in \mathbb{R}^n \times \mathbb{R}} && z = u \\
& \text{sujet à} && \nabla g_{I^*}(x^*) \tilde{d} \leq u e \\
& && \nabla h(x^*) \tilde{d} = 0
\end{aligned}$$

La solution u est clairement plus petite ou égale à $\alpha < 0$ alors que la solution du problème dual (4) est 0. On en déduit que le problème (3) n'est pas borné inférieurement et que le problème (4) n'est pas réalisable par le théorème de dualité faible. Alors il n'existe pas $\lambda \geq 0$ (μ libre et λ non tous nuls) tels que

$$\begin{aligned}
& \lambda \nabla g_{I^*}(x^*) + \mu \nabla h(x^*) = 0 \\
& \lambda \geq 0
\end{aligned} \tag{5}$$

qui est équivalent au résultat (2).

(\Leftarrow) On suppose d'abord que (2) est satisfait. Ceci veut dire qu'il n'existe pas $(\lambda, \mu) \neq 0$ tels que (5) est satisfait ou encore il n'existe pas (λ, μ) tels que (4) est réalisable. Ainsi, le dual est non réalisable. Par le théorème de dualité faible, on doit donc avoir le problème primal (3) non borné inférieurement (car réalisable en $d = 0, u = 0$), donc il existe $d \in \mathbb{R}^n, u \in \mathbb{R}$ tels que

$$\begin{aligned}
& u < 0 \\
& \nabla g_{I^*}(x^*) d \leq u e \\
& \nabla h(x^*) d = 0
\end{aligned}$$

d'où on déduit MFCQ.

□